CONSTRAINED MAXIMUM-LIKELIHOOD COVARIANCE ESTIMATION FOR TIME-VARYING SENSOR ARRAYS

David W. Rieken and Daniel R. Fuhrmann

Electronic Systems and Signals Research Laboratory
Department of Electrical Engineering
Washington University in St. Louis
St. Louis, MO 63130
rieken@essrl.wustl.edu, danf@essrl.wustl.edu

ABSTRACT

We examine the problem of maximum likelihood covariance estimation using a sensor array in which the relative positions of individual sensors change over the observation interval. The problem is cast as one of estimating a structured covariance matrix sequence. A vector space structure is imposed on such sequences, and within that vector space we define a constraint space given by the intersection of a hyperplane W_1 and the space of sequences of nonnegative definite matrices W_2. Knowledge of the changing array geometry is used to reduce the dimension of the search space. An extension of the inverse iteration algorithm of Burg et al. is proposed for finding the maximum likelihood solution.

1. INTRODUCTION

In many array signal processing applications knowledge of the observation covariance matrix is essential. Examples of such applications include MVDR beamforming and direction of arrival estimation using MUSIC. Many algorithms for estimating the covariance matrix are available. Perhaps the simplest and at the same time most common is given by

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{x}_m \mathbf{x}_m^H \tag{1}$$

which is the maximum-likelihood estimate given identical, independent, zero-mean random vectors $\mathbf{x}_m$ with covariance $\mathbf{R}$. Other estimators incorporate information about the array geometry. These are commonly called structured covariance estimators and were introduced in [1].

When the array changes shape significantly over an observation interval the statistics of the data vectors change, however. This invalidates the identical distribution assumption used to obtain (1) and the assumptions of most structured covariance estimation algorithms. The phenomenon of time-varying arrays of sensors exists in nearly all array applications. (No array is truly time-invariant, although they may be close enough to achieve the desired performance.) The effect is exaggerated, though, in towed sonar arrays which are subject to underwater currents and the maneuvering of their parent platform. An array of sensors in which each sensor is mounted on a different platform with its own propulsion also constitutes a time-varying array.

Direction-of-arrival and spectrum estimation for time-varying arrays has been addressed by a number of authors. Direction-of-arrival estimation was addressed in [2] and [3]. Fast algorithms for doing the same which are based on the eigenstructure of the matrix are presented in [4]. In [5] the EM algorithm is used to estimate the power of far-field sources using a time-varying array.

In this paper we address the problem of maximimum likelihood (ML) covariance estimation for time-varying arrays. We proceed by defining a mathematical infrastructure and applying commonly used linear algebra techniques. We then propose several search algorithms to find the covariance that maximizes the likelihood under several constraints imposed by the array motion. What results may be considered a time-varying structured covariance estimation algorithm.

2. DEFINITIONS

Let N be the number of elements in the array and $\mathbf{p}_n(t) \in R^3$ be the position of the nth element at time t. Let M denote the number of data vectors sampled by the array at times $\{t_1, t_2, \cdots, t_M\}$ with sampling frequency F_s. The mth data vector we represent by $\mathbf{x}_m$ which is a normally distributed complex random variable with mean $\mathbf{0}$ and covariance $\mathbf{R}(t_m) = \mathbf{R}_m$. The time-varying nature of the array implies that $\mathbf{R}_m$ need not equal $\mathbf{R}_{m+1}$. The problem is therefore to estimate $\mathbf{R}_m$ for $m = 1, \cdots, M$.

We make two assumptions regarding the available infor-

This research was funded by a grant from MIT Lincoln Laboratory.

mation. First, the $N \times 1$ steering vector $\mathbf{a}(\Theta, t)$ is known for all t_m and for all $\Theta \in S^2$. The nth element of the steering vector is given by

$$\mathbf{a}_n(\Theta, t) = \exp\left[-\frac{\mathbf{k}^T(\Theta)\mathbf{p}_n(t)}{\lambda}\right] \tag{2}$$

where $\mathbf{k}(\Theta)$ is the unit vector associated with the direction Θ. Secondly, the signal originating at any direction is uncorrelated with signals originating at other directions. Also, the sampling rate is such that the sampled signals are independent random variables. The time-varying covariance matrix is then given by

$$\mathbf{R}_m = \int_{S^2} \sigma^2(\Theta)\mathbf{a}(\Theta, t_m)\mathbf{a}^H(\Theta, t_m)d\Theta \tag{3}$$

where $\sigma^2(\Theta)d\Theta$ is the time-invariant power of the differential emitter at location Θ.

Since we are interested in a sequence of Hermitian matrices let us introduce the following notation:

Definition 1 *For positive integers N and M, let $V_{N,M}$ be a space such that $\bar{\mathbf{X}} \in V_{N,M}$ implies that $\bar{\mathbf{X}} \equiv [\mathbf{X}_1, \cdots, \mathbf{X}_M]$ where $\mathbf{X}_m \in C^{N\times N}$ and $\mathbf{X}_m^H = \mathbf{X}_m$.*

Observe that we denote elements of this space by capital, bold-faced letters with an overbar and the mth element of the sequence by the same letter with a subscript. For some $\alpha \in R$ and $\bar{\mathbf{X}} \in V_{N,M}$ we define scalar multiplication as

$$\alpha\bar{\mathbf{X}} \equiv [\alpha\mathbf{X}_1, \cdots, \alpha\mathbf{X}_M]. \tag{4}$$

Similarly, addition is defined element-wise,that is for $\bar{\mathbf{X}}, \bar{\mathbf{Y}} \in V_{N,M}$

$$\bar{\mathbf{X}} + \bar{\mathbf{Y}} \equiv [\mathbf{X}_1 + \mathbf{Y}_1, \cdots, \mathbf{X}_M + \mathbf{Y}_M]. \tag{5}$$

It is easy to see that under these operations $V_{N,M}$ is a vector space over R. For notational convenience we also define the following operations on vectors:

$$\bar{\mathbf{X}}\bar{\mathbf{Y}} \equiv [\mathbf{X}_1\mathbf{Y}_1, \cdots, \mathbf{X}_M\mathbf{Y}_M] \tag{6}$$

$$\bar{\mathbf{X}}^{-1} \equiv [\mathbf{X}_1^{-1}, \cdots, \mathbf{X}_M^{-1}]. \tag{7}$$

Notice that under vector addition as defined in (5) and using (6) as vector multiplication, $V_{N,M}$ forms a non-commutative ring . The multiplicative identity is then the length-M sequence of $N \times N$ identity matrices and (7) is the multiplicative inverse of $\bar{\mathbf{X}}$. With this in mind, it would be appropriate to refer to (6) and (7) as multiplication and inversion respectively.

Definition 2 $\forall\bar{\mathbf{X}}, \bar{\mathbf{Y}} \in V_{N,M}$ *let*

$$\langle\bar{\mathbf{X}}, \bar{\mathbf{Y}}\rangle \equiv \sum_{m=1}^{M} \mathrm{tr}\left(\mathbf{X}_m^H\mathbf{Y}_m\right).$$

We claim that this is an inner product on $V_{N,M}$. This is a result of the following facts which are easily proved for all $\alpha \in R$ and $\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}} \in V_{N,M}$:

$$\begin{aligned}
\langle\bar{\mathbf{X}}, \bar{\mathbf{Y}}\rangle &\in R \\
\langle\bar{\mathbf{X}} + \bar{\mathbf{Y}}, \bar{\mathbf{Z}}\rangle &= \langle\bar{\mathbf{X}}, \bar{\mathbf{Z}}\rangle + \langle\bar{\mathbf{Y}}, \bar{\mathbf{Z}}\rangle \\
\langle\alpha\bar{\mathbf{X}}, \bar{\mathbf{Y}}\rangle &= \alpha\langle\bar{\mathbf{X}}, \bar{\mathbf{Y}}\rangle \\
\langle\bar{\mathbf{X}}, \bar{\mathbf{Y}}\rangle &= \langle\bar{\mathbf{Y}}, \bar{\mathbf{X}}\rangle \\
\langle\bar{\mathbf{X}}, \bar{\mathbf{X}}\rangle &\geq 0 \text{ with equality iff } \bar{\mathbf{X}} = [\mathbf{0}, \cdots, \mathbf{0}]
\end{aligned}$$

Therefore $(V_{N,M}, \langle.,.\rangle)$ is an inner product space.

The covariance matrix sequence is an element, $\bar{\mathbf{R}}$, of $V_{N,M}$. With this in mind we can rewrite (3) as

$$\bar{\mathbf{R}} = \int_{S^2} \sigma^2(\Theta)\bar{\Psi}(\Theta)d\Theta \tag{8}$$

where

$$\bar{\Psi}(\Theta) = \{\mathbf{a}(\Theta, t_1)\mathbf{a}^H(\Theta, t_1), \cdots, \mathbf{a}(\Theta, t_M)\mathbf{a}^H(\Theta, t_M)\}.$$

The span of $\bar{\Psi}(\Theta)$ over all $\Theta \in S^2$ is a vector subspace of $V_{N,M}$. We will call this subspace W_1. It is clear from (8) that $\bar{\mathbf{R}} \in W_1$. Being a vector space, W_1 is convex and therefore path-connected. Furthermore, there exists an orthonormal basis for W_1. We will let $W_2 \in V_{N,M}$ be the space of all length-M sequences of non-negative definite Hermitian matrices. Since any covariance matrix is non-negative definite, $\bar{\mathbf{R}} \in W_2$. It can be shown that W_2 is also convex. Since the desired sequence lies within both subspaces the constraint space is their intersection $W \equiv W_1 \cap W_2$. As the intersection of two convex sets, W is also convex.

We remark that the set of matrix sequences W may not coincide exactly with the space of matrix sequences given by the model in (3), although as has been shown the latter is a subset of W. Our constraint space may contain elements outside of the cone described by (3). The discrepancy between the two spaces, and the consequences thereof, remain open questions.

We will now derive the log-likelihood function of the covariance matrix sequence for the given data set. The pdf of the data vectors is defined only for the interior of W_2:

$$f(\mathbf{x}_1, \cdots, \mathbf{x}_M) = \pi^{-NM}\left(\prod_{m=1}^{M} |\mathbf{R}_m|^{-1}\right) \times \exp\left(-\sum_{m=1}^{M} \mathbf{x}_m^H\mathbf{R}_m^{-1}\mathbf{x}_m\right). \tag{9}$$

The log-likelihood is then

$$l(\bar{\mathbf{R}}) = -\sum_{m=1}^{M} \ln|\mathbf{R}_m| - \sum_{m=1}^{M} \operatorname{tr}\left(\mathbf{x}_m^H \mathbf{R}_m^{-1} \mathbf{x}_m\right)$$

$$= -\sum_{m=1}^{M} \ln|\mathbf{R}_m| - \sum_{m=1}^{M} \operatorname{tr}\left(\mathbf{R}_m^{-1} \mathbf{S}_m\right) \qquad (10)$$

where we shall call

$$\mathbf{S}_m = \mathbf{x}_m \mathbf{x}_m^H \qquad (11)$$

the sample covariance matrix at time t_m. Observe that $\mathbf{S}_m = \mathbf{S}_m^H$ and therefore the length-M sequence of all such matrices, $\bar{\mathbf{S}}$, is an element of $V_{N,M}$. We can use the notation of definition 2 to simplify (10):

$$l(\bar{\mathbf{R}}; \bar{\mathbf{S}}) = -\sum_{m=1}^{M} \ln|\mathbf{R}_m| - \langle \bar{\mathbf{R}}^{-1}, \bar{\mathbf{S}} \rangle. \qquad (12)$$

To find the gradient of the log-likelihood function we will make use of several differentiation theorems found in [1].

Theorem 1 *For* $\mathbf{R}, \Phi \in C^{N\times N}$,

$$\frac{d}{dx} \ln|\mathbf{R} + x\Phi| = \operatorname{tr}(\mathbf{R}^{-1}\Phi).$$

Theorem 2 *For* $\mathbf{R}, \mathbf{S}, \Phi \in C^{N\times N}$,

$$\frac{d}{dx} \operatorname{tr}((\mathbf{R} + x\Phi)^{-1}\mathbf{S}) = -\operatorname{tr}(\mathbf{R}^{-1}\Phi\mathbf{R}^{-1}\mathbf{S}).$$

The directional derivative of the log-likelihood along the vector $\bar{\Phi}$ is

$$\frac{d}{dx} l(\bar{\mathbf{R}} + x\bar{\Phi}; \bar{\mathbf{S}}) = -\frac{d}{dx} \sum_{m=1}^{M} \ln|\mathbf{R}_m + x\Phi_m| + \operatorname{tr}((\mathbf{R}_m + x\Phi_m)^{-1}\mathbf{S}_m)$$

$$= -\sum_{m=1}^{M} \operatorname{tr}(\mathbf{R}_m^{-1}\Phi_m) - \operatorname{tr}(\mathbf{R}_m^{-1}\Phi_m\mathbf{R}_m^{-1}\mathbf{S}_m)$$

$$= -\sum_{m=1}^{M} \operatorname{tr}(\mathbf{R}_m^{-1}\Phi_m - \mathbf{R}_m^{-1}\mathbf{S}_m\mathbf{R}_m^{-1}\Phi_m)$$

$$= -\sum_{m=1}^{M} \operatorname{tr}\left((\mathbf{R}_m^{-1} - \mathbf{R}_m^{-1}\mathbf{S}_m\mathbf{R}_m^{-1})\Phi_m\right)$$

$$= \langle \bar{\mathbf{R}}^{-1}\bar{\mathbf{S}}\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}, \bar{\Phi} \rangle. \qquad (13)$$

Therefore the gradient of the log-likelihood function is given by

$$\nabla l(\bar{\mathbf{R}}, \bar{\mathbf{S}}) = \bar{\mathbf{R}}^{-1}\bar{\mathbf{S}}\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}. \qquad (14)$$

Note the similarity of this expression to the analogous expression for the gradient of the log-likelihood in Burg *et al.* [1]. Here the matrices have been replaced with matrix sequences.

3. ESTIMATION ALGORITHMS

One possible estimator is the projection of the sample covariance matrix sequence, $\bar{\mathbf{S}}$, onto the constraint space W. This is equivalent to selecting the point in W that is the closest to $\bar{\mathbf{S}}$ by the standard distance metric for inner product spaces:

$$d(\bar{\mathbf{X}}, \bar{\mathbf{Y}}) = \langle \bar{\mathbf{X}} - \bar{\mathbf{Y}}, \bar{\mathbf{X}} - \bar{\mathbf{Y}} \rangle^{\frac{1}{2}}. \qquad (15)$$

Because of the similarity between this estimator and classic filtering where a signal is projected onto the subspace of all signals which satisfy a certain constraint, we will refer to this estimator as the sample sequence filter. Since W is the intersection of two convex spaces we employ the method of projection onto convex sets (POCS) in which the estimate is first projected onto W_1 and then onto W_2. This iterative procedure continues until the improvement in likelihood gained with each iteration is negligible. Since W_1 is a vector subspace the projection of a vector $\bar{\mathbf{X}}$ onto W_1 is given by

$$\bar{\mathbf{X}}' = \sum_{l=1}^{L} \langle \bar{\mathbf{X}}, \bar{\Phi}_l \rangle \bar{\Phi}_l \qquad (16)$$

where $\bar{\Phi}_l$ are the members of an orthonormal basis of W_1 and L is the dimension.

Projection onto W_2 for the given inner product is found in [6]. First, the eigendecomposition is determined:

$$\mathbf{X}_m = \Gamma_m \Lambda_m \Gamma_m^{-1} \qquad (17)$$

Then the projection onto the set of non-negative definite matrices is given by setting the negative eigenvalues to 0:

$$\mathbf{X}'_m = \Gamma_m \max(\Lambda_m, \mathbf{0}) \Gamma_m^{-1}. \qquad (18)$$

The projection of the sequence $\bar{\mathbf{X}}$ onto W_2 is the element-wise projection of each $\mathbf{X}_m$ as described by this equation.

The sample sequence filter, by its definition, finds the sequence which is in the constraint space and the closest to the sample sequence by the distance metric given in (15). Experience has shown, however, that the best estimate is rarely the closest to the sample sequence. We therefore propose searching the constraint space for the maximum likelihood estimate using the filtered sample sequence as a starting point.

Each of the search algorithms which we will consider proceed by calculating a search direction, $\bar{\mathbf{D}} \in W_1$, along

which the likelihood function must be maximized. That is, in each iteration we determine a $\bar{\mathbf{D}}$ and then find λ_0 such that

$$\lambda_0 = \arg\max l(\hat{\bar{\mathbf{R}}}^{(i)} + \lambda \bar{\mathbf{D}}).$$

The updated estimate is

$$\hat{\bar{\mathbf{R}}}^{(i+1)} = \hat{\bar{\mathbf{R}}}^{(i)} + \lambda_0 \bar{\mathbf{D}}.$$

This iterative process should be allowed to continue until the gradient is sufficiently close to being orthogonal to W_1, that is until

$$\frac{\sum_{i=1}^{L} \langle \nabla l(\hat{\bar{\mathbf{R}}}; \bar{\mathbf{S}}), \bar{\Phi}_i \rangle^2}{\langle \nabla l(\hat{\bar{\mathbf{R}}}; \bar{\mathbf{S}}), \nabla l(\hat{\bar{\mathbf{R}}}; \bar{\mathbf{S}}) \rangle} < \epsilon. \tag{19}$$

Observe that since l is defined only on the interior of W_2 and $\bar{\mathbf{D}} \in W_1$, the estimate will be within the constraint space at each iteration.

Perhaps the most obvious approach to calculating $\bar{\mathbf{D}}$ is to use the gradient in (14). $\bar{\mathbf{D}}$ can be the projection of the gradient onto W_1. Alternatively, a conjugate gradient direction can be calculated by incorporating memory of previous search directions. We suggest, however, a modification of the inverse iteration algorithm proposed by Burg, *et al.* in [1]. Burg's algorithm was designed for estimation of a single matrix rather than a sequence of matrices but is easily generalized for sequences. For some estimate $\hat{\bar{\mathbf{R}}}^{(i)}$ select a search direction, $\bar{\mathbf{D}}$, such that $\nabla l(\hat{\bar{\mathbf{R}}}^{(i)}; \bar{\mathbf{S}} - \bar{\mathbf{D}})$ is orthogonal to W_1. Clearly, if $\bar{\mathbf{D}} = 0$ then $\hat{\bar{\mathbf{R}}}^{(i)}$ is the maximum likelihood estimate since the likelihood gradient is orthogonal to the constraint space at that point.

Before the modified inverse iteration algorithm may be seriously considered, though, one must ask whether a stable point of the algorithm maximizes the likelihood function within the constraint space. That is, does each iteration of the algorithm lead to an improvement in likelihood for a nonzero search direction? The answer is yes, as shown by the following theorem:

Theorem 3 *Suppose there exists $\bar{\mathbf{D}} \neq \mathbf{0} \in W_1$ such that $\nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}} - \bar{\mathbf{D}})$ is orthogonal to W_1. Then there exists $\lambda \in R$, $\lambda \neq 0$, such that $l(\bar{\mathbf{R}} + \lambda \bar{\mathbf{D}}; \bar{\mathbf{S}}) > l(\bar{\mathbf{R}}; \bar{\mathbf{S}})$.*

Proof: By way of contradiction, suppose that

$$\arg\max l(\bar{\mathbf{R}} + \lambda \bar{\mathbf{D}}; \bar{\mathbf{S}}) = 0.$$

This implies that $\nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}})$ is orthogonal to $\bar{\mathbf{D}}$. That is,

$$\langle \nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}}), \bar{\mathbf{D}} \rangle = 0.$$

Therefore,

$$\begin{aligned} \langle \nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}} - \bar{\mathbf{D}}), \bar{\mathbf{D}} \rangle &= \langle \bar{\mathbf{R}}^{-1}(\bar{\mathbf{S}} - \bar{\mathbf{D}})\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}, \bar{\mathbf{D}} \rangle \\ &= \langle \nabla l(\bar{\mathbf{R}}, \bar{\mathbf{S}}) - \bar{\mathbf{R}}^{-1}\bar{\mathbf{D}}\bar{\mathbf{R}}^{-1}, \bar{\mathbf{D}} \rangle \\ &= -\langle \bar{\mathbf{R}}^{-1}\bar{\mathbf{D}}\bar{\mathbf{R}}^{-1}, \bar{\mathbf{D}} \rangle \end{aligned}$$

Since $\bar{\mathbf{D}} \in W_1$ we know that $\langle \nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}} - \bar{\mathbf{D}}), \bar{\mathbf{D}} \rangle = 0$. Therefore

$$\langle \bar{\mathbf{R}}^{-1}\bar{\mathbf{D}}\bar{\mathbf{R}}^{-1}, \bar{\mathbf{D}} \rangle = 0.$$

It can be shown that this implies that $\mathbf{R}_m^{-1}\mathbf{D}_m = \mathbf{0}$ for all m. Therefore $\mathbf{D}_m = \mathbf{0}$ for all m which is a contradiction. ■

We now concentrate on finding the direction which satisfies the necessary condition on the gradient. This is equivalent to finding $\bar{\mathbf{D}}$ which satisfies

$$\langle \bar{\mathbf{R}}^{-1}(\bar{\mathbf{S}} - \bar{\mathbf{D}})\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}, \bar{\Phi}_i \rangle = 0 \tag{20}$$

for all i. We note that this is a system of equations which are linear in $\bar{\mathbf{D}}$ and that therefore a closed-form solution exists. Since $\bar{\mathbf{D}} \in W_1$ there exist real α_j such that

$$\bar{\mathbf{D}} = \sum_j \alpha_j \bar{\Phi}_j. \tag{21}$$

Substituting (21) into (20) and rearranging we get

$$\begin{aligned} &\langle \bar{\mathbf{R}}^{-1}(\bar{\mathbf{S}} - \bar{\mathbf{D}})\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}, \bar{\Phi}_i \rangle \\ &\quad = \langle \bar{\mathbf{R}}^{-1}(\bar{\mathbf{S}} - \sum_j \alpha_j \bar{\Phi}_j)\bar{\mathbf{R}}^{-1} - \bar{\mathbf{R}}^{-1}, \bar{\Phi}_i \rangle \\ &\quad = \langle \nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}}), \bar{\Phi}_i \rangle - \sum_j \alpha_j \langle \bar{\mathbf{R}}^{-1}\bar{\Phi}_j\bar{\mathbf{R}}^{-1}, \bar{\Phi}_i \rangle. \end{aligned} \tag{22}$$

Therefore we need only find $\alpha \in R^L$ such that

$$\mathbf{A}\alpha = \mathbf{B} \tag{23}$$

where

$$\mathbf{A}_{ij} = \langle \bar{\mathbf{R}}^{-1}\bar{\Phi}_j\bar{\mathbf{R}}^{-1}, \bar{\Phi}_i \rangle \tag{24}$$

$$\mathbf{B}_i = \langle \nabla l(\bar{\mathbf{R}}; \bar{\mathbf{S}}), \bar{\Phi}_i \rangle. \tag{25}$$

4. COMPUTER SIMULATION

We have simulated a uniform linear array (ULA) consisting of $N = 5$ isotropic sensors which is rotating with rotational velocity ω about the center element. The axis of rotation is orthogonal to the axis of the ULA. There are 3 source fields impinging upon the array which originate at (azimuth, elevation) = $(45°, 0°)$, $(85°, 20°)$, and $(110°, 0°)$. Here azimuth is the the angle made with the axis of the array at $t = 0$ and within the plane of rotation. Elevation

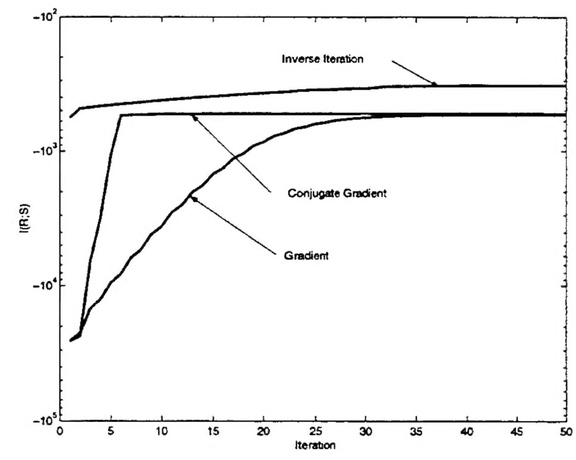

Figure 1: Algorithm convergence rate comparison.

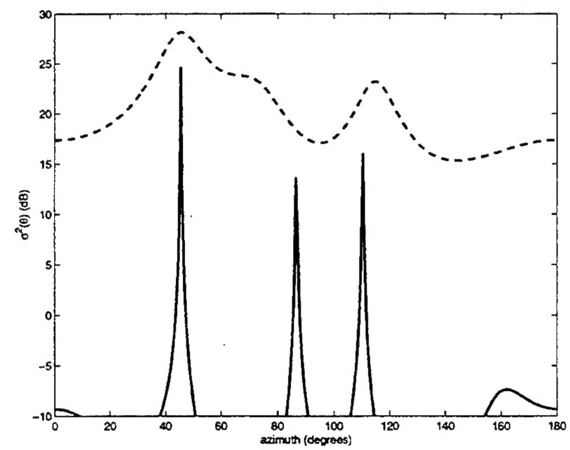

Figure 2: MVDR spectrum estimated from the first matrix in the sequence. The dashed line is that obtained from the sample sequence filtering procedure. The solid line was calculated from the ML sequence.

is the angle made with the plane of rotation. For example, $(90°, 10°)$ would describe a direction orthogonal to the initial array axis and $10°$ above the plane of rotation. Each of the sources are assumed to be narrowband with wavelength λ and the separation between elements in the ULA is $\frac{\lambda}{2}$. The power of each source at the array is 30dB, 15dB, and 20dB respectively. Receiver noise is 0dB. The rotational velocity of the array is $\omega = 2\pi$ rad/sec, the sampling rate is $F_s = 32s^{-1}$, and the number of samples collected is $M = 16$. Therefore, the array gathers 16 data vectors while completing a half rotation. Since the statistics of the data vector change dramatically over this observation interval one expects the covariance estimator in (1) to perform badly. It is unclear even what steering vector to use with this covariance estimate in, for example, an MVDR or MUSIC estimator. That makes this scenario a good candidate for covariance matrix *sequence* estimation.

Each of the algorithms considered begins by calculating the sample covariance matrix sequence and filtering it by the method of POCS using the projections in (16) and (18). The $\bar{\Phi}_l$ are obtained by Gram-Schmidt orthonormalization of the set of vectors $\Psi(\Theta)$ where Θ is a discretization of the 2-sphere. For this example, the dimension of W_1 is $L = 76$. Note that the dimension of $V_{N,M}$ is $N^2M = 400$ and we have managed to eliminate 324 parameters of the matrix sequence by applying information about the motion of the array.

The ML search routine follows the filtering. We applied a gradient and conjugate gradient search algorithm in addition to the modified inverse iteration algorithm. The likelihood of the estimate at each iteration is plotted in Figure 1 for each of the algorithms.

Observe that the gradient and conjugate gradient algorithms converge to points with the same likelihood. The conjugate gradient reaches this point in fewer iterations, which is to be expected. However, the convergence point of the inverse iteration algorithms exceeds the likelihood of the estimate obtained from either gradient algorithm after only a few iterations. Inspection of the likelihood gradient at what appears to be the convergence point of the gradient algorithms reveals that it is not orthogonal to W_1 and that while successive iterations yield only slight improvement in likelihood, they have failed to reach a local maxima. One possibility is that they have stumbled upon a "ridge" in the likelihood function. It is clear that, in this example at least, the inverse iteration algorithm reaches a solution in fewer iterations than even the conjugate gradient algorithm. It should be noted, however, that finding the solution to (23) requires more compuation than calculating the likelihood gradient and projecting it onto W_1.

To demonstrate the validity of the ML estimate, the MVDR spectrum corresponding to the first matrix in the sequence has been calculated and plotted in Figure 2. The spectrum is calculated using

$$\hat{\sigma}^2(\theta) = \frac{1}{\mathbf{a}^H(\theta, t_1)\hat{\mathbf{R}}_1^{-1}\mathbf{a}(\theta, t_1)}. \tag{26}$$

$\hat{\mathbf{R}}_1$ is the first matrix in the sequence obtained by the inverse iteration algorithm since the other two algorithms failed to produce a maximum likelihood estimate. The position of each of the sources is easily ascertained from the plot as is a feeling for their intensities. Also plotted is the spectrum obtained from just the sample sequence filter. While peaks which correspond to two of the sources can be seen, the third is lost and the background noise is quite high. This

demonstrates the necessity of the ML search algorithms.

5. CONCLUSIONS

We have developed an algorithm for estimating the sequence of matrices which are the covariances of the data vectors of a time-varying sensor array. Since each matrix in the sequence is structured to satisfy (3) this may be classified as a structured covariance estimation algorithm in which the sequence of matrices itself is structured. This method will have good performance for time-varying arrays in which the array motion is periodic, as with the rotating ULA, since the constraint space basis need not be continuously recalculated, which relieves the processor of some of the computational burden. It has been demonstrated that the modified inverse iterations algorithm can converge faster and more reliably than a simple gradient search algorithm, although with increased computational complexity per iteration.

6. REFERENCES

[1] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, pp. 963–974, Sept. 1982.

[2] A. Zeira and B. Friedlander, "Direction finding with time-varying arrays," *IEEE Trans. on Signal Processing*, vol. 43, pp. 927–937, April 1995.

[3] J. Sheinvald, M. Wax, and A. J. Weiss, "Localization of multiple sources with moving arrays," in *Proc. ICASSP '97*, pp. 3521–3524, 1997.

[4] B. Friedlander and A. Zeira, "Eigenstructure-based algorithms for direction finding with time-varying arrays," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 32, pp. 689–701, April 1996.

[5] D. W. Rieken, A. D. Lanterman, and D. R. Fuhrmann, "Spatial spectrum estimation for time-varying arrays using the EM algorithm," in *Proc. of the 38th Ann. Allerton Conf. on Communication, Control, and Computing*, October 2000.

[6] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra and Its Applications*, vol. 103, pp. 103–188, May 1988.